DOCUMENT RESUME

ED 449 213 TM 032 358

AUTHOR Weber, Deborah A.
TITLE Restriction of Range: The Truth about Consequences and Corrections.
PUB DATE 2001-02-00
NOTE 16p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, February 1-3, 2001).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Correlation; Reliability; Scores; Validity
IDENTIFIERS Power (Statistics); *Restriction of Range

ABSTRACT

Restriction of range occurs whenever design or circumstances abbreviate the values of one or both variables being correlated. Range restriction has an impact on score validity, score reliability, and statistical power. As a result, researchers need a more sophisticated concept of the relationship between variance and correlation than the simple assumption that when variance goes down, correlations decrease. The shape of the distribution and selection procedures also influence whether restriction of range increases, decreases, or does not affect correlation. (Contains 16 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Running head: RESTRICTION OF RANGE: THE TRUTI

ED 449 213

Restriction of Range:

The Truth About Consequences and Corrections

Deborah A. Weber

Texas A&M University 77843-4225

D. A. Weber

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM032358

Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, February 1-3, 2001.

 BEST COPY AVAILABLE

Abstract

Range restriction impacts score validity, score reliability, and statistical power. Thus, researchers need a more sophisticated concept of the relationship between variance and correlation than simply assuming that when variance goes down, correlations decrease. The shape of the distribution and selection procedures also influence whether restriction of range increases, decreases, or does not affect correlation.

Restriction of Range:

The Truth About Consequences and Corrections

Like parrots, many of us graduate students are conditioned to respond that if variance decreases then correlation decreases. However, the relationship between variance and correlation depends on "the way in which changes in variability are brought about" (Huck, 1992, p. 258). When the range of values of one or both variables being correlated is curtailed (Gall, Borg, & Gall, 1996), then the resulting Pearson r may be larger, smaller, or equal to the Pearson r of the complete data set. The relationships between the variables and the selection procedures determine the influence of restriction of range on correlation, score validity, score reliability, and statistical power.

Types of Selection

Restriction of range occurs whenever design or circumstances abbreviate the values of one or both variables being correlated. Crocker and Algina (1986) distinguished three types of selection: (a) natural attrition, (b) explicit selection, and (c) incidental selection.

Natural attrition can account for restriction of range "whenever subjects at the high or low end of the criterion continuum tend to leave the setting before criterion data can be collected" (Crocker & Algina, 1986, p. 227). For example, a study on Downs syndrome babies at age two may result in natural selection because many of the more severely effected babies may not survive to the age at which the data are being obtained.

Explicit selection occurs when "the test being validated is used for selection purposes before its validity has been established" (Crocker & Algina, 1986, p. 227); therefore, participants are intentionally excluded because their score falls below a certain point. According to a study by Duan and Dunlap (1997), when restriction of range occurs due to explicit selection, corrected

correlations provide estimates of the population parameters that are more accurate. However, if the selection ratio is small, a correlation that has been corrected is not an accurate estimate of the population r. Standard error tends to increase as the ratio decreases and contrarily the standard error decreases as the sample/population size increases.

Incidental selection occurs when selection "is made on the basis of some other variable that is correlated with the predictor test being validated" (Crocker & Algina, 1986, p. 226). For example, when validating a college admissions test, students are admitted regardless of their scores on the admissions test due to their high school grades. A ceiling (many examinees have very high scores) or floor (many have very low scores) effect may also cause restriction of range (Crocker & Algina, 1986; Walsh, 1996).

While restriction of range may be either direct or indirect, sampling variables that are homogeneous cause underestimation of the degree of the relationship between variables (Walsh, 1996). When sampling actions are biased in relation to one of the two variables a direct restriction of range occurs. Indirect restriction of range occurs when a third or extraneous variable is restricted.

Effect of Restriction of Range on Score Validity

When a test is used to select people who promise to be successful in college or in the workplace, the test scores are used in conjunction with the criterion variable to determine how accurately the scores predict behavior. Prior to determining who is accepted and who is rejected, the scores of the complete population of applicants are available. While the population of applicants may not be representative of society, the validity statistics from the complete pool of test-takers are the "validity statistics in which we are basically interested" (Thorndike, 1949, p.

170). However, in practice, one rarely has data for the total applicant population because individuals who are likely to be accepted for training or employment tend to be selected.

Once a selection procedure has occurred to differentiate certain individuals who are applying the criterion scores of the selected applicants cease to be representative of the group of applicants. After selection, criterion scores are only available for the applicants who are accepted, and then only for the applicants who are not lost through attrition (Crocker & Algina, 1986). As Thorndike (1949) noted, "as soon as some selection procedure operates to pick certain type of individuals from among those applying, the group for whom criterion records subsequently become available ceases to be representative of the general group of applicants" (p. 170).

The problem becomes estimating the correlation for the full applicant population using an incomplete sample (Gross & Kagen, 1983). As Crocker and Algina (1986) noted,

> Because the most common form of validity evidence is the correlation coefficient between predictor and criterion scores, it is important to recognize the restriction of the range of scores on either the predictor or criterion many result in attenuation of the observed validity coefficient. (p. 226)

Depending on the relationship between the predictor and criterion variables and/or selection procedures, restriction of range causes the correlation coefficient of the selected group to increase, decrease, or remain the same as that in the complete group (Huck, 1992).

Effect of Restriction of Range on Correlation

As stated by Walsh (1996), Pearson r is a coefficient of correlation that indicates the direction and degree of association between two variables. Additionally, coefficient r evaluates only the linear relationship between two observed (weight, height) or measured (self-concept) variables.

two latent (regression) or synthetic (e scores) variables, or between one observed and one synthetic variable. Pearson r is a vital part of calculations in parametric statistics and in measurement and it is imperative that its nature and factors affecting Pearson r are understood.

An article written by Huck in 1992 posed the question "What is the connection between Pearson's r and variability" (p. 258)? According to Huck, the connection between heterogeneity and Pearson's r is not easy to predict and researchers are not in agreement regarding the connection between the magnitude of Pearson's r and variability. For example, which response would be considered the correct response to the question Huck posed? Huck (1992) suggested that "The connection between Pearson's r and variability is (a) as variability increases, r increases, (b) as variability decreases, r decreases, or (c) changes in variability neither increase or decrease r" (p. 253).

Various ideas regarding variability and Pearson's r exist due to statements in textbooks and by instructors which indicate each of the three above incongruous answers is the correct answer to the question poised. How can apparently contradictory statements be written to describe accurately the connection between variability and Pearson's r? Very few individuals possess the knowledge to recognize that none of the above answers is better than the other. The correct answer is, "it depends" (Huck, 1992, p. 254). According to Huck three factors affect score variability including linear transformations of data, and errors of measurement, and restriction of range.

Statements in textbooks lead students to believe that score variability and r are directly correlated. According to Huck (1992) textbooks often graphically illustrate the relationship between restriction of range and correlation by indicating that restriction of range always brings about a decrease in the variability of scores. These graphic illustrations correctly direct one to

conclude that a decrease in the variance also brings about a decrease in the correlation coefficient but do not give sufficient attention to exceptions. In addition to pictures that illustrate restriction of range, textbooks often use statements that make the same point. For example as cited in Huck (1992, p. 225),

> If, in a study of the relationship between measures of two traits, we selected two groups of individuals such that one group showed greater variability in these measures than the other, we would find that the coefficient of correlation r between the measures would be greater for the more variable than for the more homogeneous group. (Lindquist,1942, pp. 195-196)

Or, according to Glass and Hopkins (1984, p. 92) "Other things being equal, *the greater the variability among the observations, the greater the value of r*". Such statements can only lead a reader to conclude that when variance goes down, correlation decreases. The problem is that these dynamics are <u>not</u> universally true.

Unfortunately, the above posed question does not have an easy answer. The answer "it depends" reflects the fact that changes in the magnitude or <u>r</u> and as a function of variability depend upon the nature of the changes in variability. Restriction of range <u>may</u> be due to the removal of data points that contain X or Y scores that lie below (or above) a mid-range cut-off value (Huck, 1992).

But, when the relationship between the predictor and criterion variables is not linear, "restriction of range could serve to increase (not decrease) the correlation" (Huck, 1992, p. 260). The correlation also increases when values in the middle of the range of scores are selected to be discarded.

If the relationship between the two variables is perfectly linear, restriction of range will not affect the correlation. As Lord and Novick (1968) noted, "It is well known that the size of a correlation coefficient depends very much on the nature of the population in which measurements are made" (p. 129).

Assuming that restriction of range always decreases the correlation between the predictor and criterion variable is rash. Selection procedures and/or the nature of the relationship between the predictor and criterion variables determine the effect of restriction of range on correlation.

Effect of Restriction of Range

Power

Restriction of range can also decrease statistical power (Hallahan & Rosenthal, 1996). Statistical power is "the probability of rejecting the null hypothesis when it is false" (Hinkle, Wiersma & Jurs, 1998, p. 620). Hinkle et al. (1998) presented four factors that determine statistical power: sample size, level of statistical significance, directional nature of a test, and effect size. Hallahan and Rosenthal (1996) concurred that statistical power is influenced by any factor "that has implications for any of these parameters" (p. 495).

Score Reliability and Validity

Through selection, restriction of range decreases the variance on one or both variables in a bivariate correlation, consequently also affecting score reliability and score validity. Thus, Karl Pearson (1903) long ago wrote "We must always bear in mind this all-important fundamental conception, that natural or artificial selection, or even random sampling, are in themselves active factors in the modification (i.e., creation destruction, or reversal) of correlation" (p. 29). Therefore, the nature of selection must be accounted for when addressing restriction of range.

According to Pearson correlation is completely dependent upon selection which can create, destroy, or reverse correlation.

According to Thompson and Vacha-Haase (2000) "measurement textbooks on the whole do poorly at accurately communicating fundamental measurement concepts to our graduate students, at least as regards score reliability" (p. 175). Additionally,

> The fact that homogeneity or the heterogeneity of the scores of the people being measured directly affects score reliability obviously means that reliability is not indelibly and unalterably stamped into a test booklets during the printing process. Score reliability changes as a measure is administered to different samples, partially as a function of the score standard deviation in a given group. (Thompson & Vacha-Haase, 2000, p. 177)

Correcting for Restriction of Range

In 1903, Pearson published formulas that correct for restriction of range when the variables being studied originate from a normally distributed population. In 1949, Thorndike identified three separate cases involving restriction of range "differentiated in terms of the variables that has served as the instrument for curtailment and in terms of the variable whose standard deviation in the unselected population is known" (p. 172). Responsible application of the formulas for correcting for restriction of range depends upon accurately determining the nature of the range restrictions.

Crocker and Algina (1986) cited procedures for estimating validity of coefficients discussed by Thorndike, Gulliksen, and Lord and Novick. These procedures however, require assumptions "that the linear regression of Y (criterion variable on X (predictor) is the same over all values of X, that is, for the selected and the unselected group" (Crocker & Algina, 1986, p. 227).

Additionally, it must be assumed that $O^2_{y.x}$ is the same of all X values. Errors may occur in correcting validity coefficients if these assumptions are violated. Instead of relying on statistical corrections, it is better to avoid restriction of variance in a validation study (Crocker & Algina, 1986).

Assumptions of Linearity and Homoscedasticity

As stated in Thompson and Vacha-Haase (2000), the degree of the scores homogeneity affects consistency. According to Cunningham (1986, p. 114)

> When scores are bunched together, a small [random measurement error] change in raw score will lead to large changes in relative position. If scores are spread out (variability is high), it is more likely that the relative position in the group will remain stable across the two forms of the test and the correlation coefficient will be relatively large.

Therefore, if there is a large difference in scores, the effects of the scores shifting positions is less likely.

Corrections for restriction of range assume the linearity of regression and homoscedasticity (Holmes, 1990, p. 19). Simply, to accurately correct for restriction of range: (a) the two variables must have a linear relationship throughout the range of values (Gall et al., 1986); and (b) "the standard deviations of all conditional distributions...[must be] equal" (Hinkle et al., 1998, p. 618). Violating the assumptions of linearity and homoscedasticity compromises the accuracy of the corrected correlation of the predictor and criterion variables for the applicant population.

Huck (1992) illustrated that restricting range of value in a non-linear distribution may actually increase the correlation coefficient of the restricted group. Generally, if the population correlation is small, selection procedures extreme, and the selected group is small, "the greater

the advantage of the uncorrected correlation over the corrected correlation" (Gross & Kagen, 1983, p. 393). When the assumption of homoscedasticity is violated, "the usual correction will generally be inadequate" (Holmes, 1999, p. 22).

According to Gross and Kagen (1983), when validating a test used for selection purposes (x) it is common to deal with incomplete data and therefore complete test-criterion (y) scores are not typically available. With an incomplete data set of xy, an estimation of the correlation in the population can be made using either the uncorrected or corrected correlation values. Small sample sizes, low population correlation, and extreme selection often leads to a corrected correlation value that has a considerably smaller expected square mean than the uncorrected correlation value. However, an uncorrected value is always more biased than the corrected value.

Restriction of range refers to the statistical problem of estimating the xy correlation (ρ_{xy}) for the entire population using an xy sample that is incomplete (Gross & Kagen, 1983). When estimating the complete population being studied two choices are available: either compute the xy correlation using the available xy data or use the standard correction formula. When using the available xy data (Gross & Kagen, 1983, p. 390),

$$r_a = s_{xya}/s_{xa}s_{ya}$$

where s_{xya} = xy covariance, s_{xa} = standard deviation of x, and s_{ya} = standard deviation of y. In the standard correction formula

$$r_c = r/[r_a^2 + (s_{xa}^2/s_{xt}^2)(1 - r_a^2)]^{1/2}$$

where s_{xa}^2 = variance of x in the selected group Na and S_{xt}^2 = variance of x in the total group N. The absolute value of r_c will characteristically exceed r_a.

Which estimate is the better estimate to use when dealing with restriction of range? The common assumption is that r_c is the superior estimate due to the inequality of the relationship

between r_c and r_a. Additionally, the bias of r_a will be greater than the bias of r_c. Because r_c is the highest estimator or population correlation it will be unbiased in large samples.

However, r_c and r_a can also be compared not only in terms of estimates of population correlation bias but also using the expected mean square error (EMSE). According to Gross and Kagen "It can be argued that EMSE represents a more meaningful criterion of the accuracy of estimation than the criterion of bias" (1983, p. 390). The EMSE can be computed as:

$$EMSE = BIAS^2 + SAMPLING\ VARIANCE$$

due to the reflection of both bias of the estimation and sampling variance by the EMSE criterion. This method compensates for the possible bias of one estimator due to a smaller EMSE (as the result of difference in sampling variance). Additionally, this method is a better choice when sample size is small due to the inverse relationship between sample size and sample variance.

When using the EMSE, in small samples r_a will likely to be more biased than r_c but r_a may have a substantially smaller EMSE value than r_c. According to Gross and Kagen (1983) "it can be advantageous to estimate the population correlation using the uncorrected correlation (r_a) rather than the corrected correlation (r_c)" (p. 391). Thus, not correction for restriction of range (using r_a instead of r_c) is most beneficial when the group selected is small in size, the p_{xy} (xy relationship) is low, and selection is extreme.

While r_a is the estimate of choice for small group sizes, it becomes less so when extreme selection measures are not used and as the p_{xy} (xy relationship) increases. When the population is not normal or the selection is not a function of x it is unclear which estimate to use. Unfortunately, current research has focused on large sample sizes in which EMSE was unable to be computed due to lack of reporting or variance of the estimates (Gross & Kagen, 1983).

According to Duan and Dunlap (1997), if "r is obtained from a random sample of the population" (p. 254) then Pearson r is an efficient estimate of the population r. In reality however, it is often difficult to meet this condition. For example, a study using Graduate Record Exam (GRE) scores to study predictive validity is unable to meet this condition due to the impossibility of obtaining scores on the criterion variable for students not admitted. In this situation restriction of range affects validity coefficients by underestimating the population r. Given this condition, correcting for restriction of range is recommended (Duan & Dunlap, 1997).

As stated by Duan and Dunlap (1997), even though Thorndike was often credited for bringing restriction of range formulas into general use, it was Pearson who first devised these formulas in 1903. Pearson's equations were based on the following assumptions: "(a) linearity of regression of Y on X; (b) homoscedasticity of the error distribution; and (c) bivariate normality" (Duan & Dunlap, 1997, p. 255).

Conclusion

Because most parametric tests involve examining the relationships between two variable, the understanding of factors that affect correlation is critical (Walsh, 1996). It can be quite common for textbooks and instructors to make statements that are often contradictory. Therefore, when seemingly easy answers to questions are in fact quite complicated, it is important that an individual possess the knowledge and wisdom to distinguish the correct from incorrect choices.

References

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Orlando, FL: Harcourt Brace Jovanovich.

Cunningham, G. K. (1986). Educational and psychological measurement. New York: Macmillan.

Duan, B., & Dunlap, W. P. (1997). The accuracy of different methods for estimating the standard error of correlations corrected for range restriction. Educational and Psychological Measurement, 57, 245-265.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). Educational research: an introduction (6th ed.). White Plains, NY: Longman Publishers USA.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Gross, A. L., & Kagen, E. (1983). Not correcting for restriction of range can be advantageous. Educational and Psychological Measurement, 43, 389-396.

Hallahan, M., & Rosenthal, R. (1996). Statistical power: concepts, procedures, and applications. Behavior research and therapy, 34, 489-499.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1998). Applied statistics for the behavior sciences (4th ed.). Boston: Houghton Mifflin.

Holmes, D. J. (1990). The robustness of the usual correction for restriction in range due to explicit selection. Psychometrika, 55, 19-32.

Huck, S. W. (1992). Group heterogeneity and Pearson's r. Educational and Psychological Measurement, 52, 253-260.

Lindquist, E .F. (1942). A first course in statistics. Cambridge, MA: Riverside.

Lord, F. M., & Novick, R.M. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Pearson, K. (1903). Mathematical contributions to the theory of evolution: XI. On the influences of natural selection of the variability and correlation of organs. Transactions of the Royal Society, London, Series A, 200, 1-66.

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. Educational and Psychological Measurement, 60, 174-195.

Thorndike, R. L. (1949). Personnel selection. New York: John Wiley.

Walsh, B. D. (1996). A note of factors that attenuate the correlation coefficient and its analogs. In B. Thompson (Ed.), Advances in social science methodology (Vol. 4, pp. 21-32). Greenwich, CT: JAI Press.